Analysis of a Gross Counting Decision Metric for use in Threat Detection During Cargo Container Inspection

James Hall

Lawrence Livermore National Laboratory

28 April 2006

Nuclear Car Wash (NCW) Project Report

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

Auspices Statement

This work was performed under the auspices of the U. S. Department of Energy (DOE) by the University of California, Lawrence Livermore National Laboratory (LLNL) under Contract No. W-7405-Eng-48.

Analysis of a gross counting decision metric for use in threat detection during cargo container inspection

James Hall

Lawrence Livermore National Laboratory, Livermore, CA 94550

Abstract: LLNL is actively engaged in the development of a variety of advanced technologies for use in detecting potential threats in sea-going cargo containers, particularly the presence of hidden special nuclear materials (SNM). One such project is the so-called "Nuclear Car Wash" (NCW), which uses a high-energy neutron probe to scan the container. High-energy, β-delayed γ rays emitted during the decay of short-lived, neutron-induced fission products are then taken as a signature of fissionable material. There are a number of different threat decision metrics that one could imagine using in conjunction with an inspection system such as the NCW; however, the most straightforward approach might be to simply compare the total number of counts that our detector records during some suitably chosen time interval to the average background signal that one would expect from a "clean" container during the same interval. The purpose of this report is to describe the basic statistical properties of a decision metric of this sort and outline the procedures for using it in experimental practice.

1. Introduction

LLNL is actively engaged in the development of a variety of advanced technologies for use in detecting potential threats in sea-going cargo containers, particularly the presence of hidden special nuclear materials (SNM). One such project is the so-called "Nuclear Car Wash" (NCW), which uses a high-energy ($E_n \approx$ 3.5-7.0 MeV), D(d,n)^{3}He neutron probe to scan the container [1]. High-energy ($E_\gamma \geq$ 2.5 MeV), β-delayed γ rays emitted during the decay of short-lived, neutron-induced fission products may then be taken as a signature of fissionable material [2] since there are relatively few neutron-induced or naturally-occurring interference signals above a few MeV in energy. The ultimate sensitivity of this (or any other) cargo inspection system will depend on the acceptable false alarm rate (P_{fa}), the required probability of detecting *real* threats (P_d) and, of course, the decision metric used to make the call.

There are a number of different, potentially very powerful, threat decision metrics that one could imagine using in conjunction with an inspection system such as the NCW (*e.g.* those recently proposed by Pruet, *et al*. [3], which involve multi-parameter fits to the time structure of the observed data); however, the most straightforward approach (and perhaps our initial "call to arms") might be to simply compare the total number of counts that our detector records during some suitably chosen time interval (*e.g.* one which will allow us to capture the majority of the delayed γ rays from short-lived SNM fission products, if any) to the average background signal that one would expect from a "clean" container during the same interval. The purpose of this report is to describe the basic statistical properties of a decision metric of this sort and outline the procedures for using it in experimental practice.[1]

---

[1] Note that, while this may on the surface seem hardly worthy of a formal report, the devil lies in the details and the correct procedures for using a metric of this sort may not be what one would naively expect …

2. Analysis in terms of measured counts

In order to make semi-quantitative estimates for critical threat signal levels and their associated signal-to-noise ratios, we will assume that the total number of counts S'' that our detector records during a time interval Δt is a discrete random variable drawn from a Gaussian-like distribution whose mean and variance are given by:

$$\overline{S''} = \langle S'' \rangle = \langle T'' + B'' \rangle = \overline{T''} + \overline{B''} \tag{1a}$$

$$\sigma_{S''}^2 = \sigma_{T''}^2 + \sigma_{B''}^2 = \overline{T''} + \sigma_{B''}^2 \; , \tag{1b}$$

where $T'' \geq 0$ is our SNM threat signal (assumed to be distributed Poisson), $\overline{T''}$ is the hypothetical mean value of T'' (*e.g.* if we could sit and take repeated measurements of S''), $B'' > 0$ is the actual background signal present during the interval and $\overline{B''}$ is the mean value of B'' (averaged over *many* Δt intervals and *many* containers). The distribution of the background signal is still a subject of debate at this point. If B'' is distributed Poisson (as is likely the case, since the background will be due largely to neutron-induced and naturally-occurring radiation sources), then $\sigma_{B''}^2$ will be $\approx \overline{B''}$; however, there is also anecdotal evidence to suggest that $\sigma_{B''}^2$ may be as large as $\overline{B''}^2$ (pessimistic case) [4]. Since the vast majority of containers scanned will (we hope) *not* contain SNM (*i.e.* $T'' = 0 \Rightarrow S'' = B''$), this issue will resolve itself in time as we inspect more and more containers with a variety of cargos and eventually establish stable values for $\overline{B''}$ and $\sigma_{B''}^2$ using standard unbiased estimators such as:

$$\hat{\overline{B''}} = \sum_{i=1}^{N} B_i'' / N \tag{2a}$$

$$\hat{\sigma}_{B''}^2 = \sum_{i=1}^{N} \left(B_i'' - \overline{B''} \right)^2 / (N-1) \approx \sum_{i=1}^{N} \left(B_i'' - \overline{B''} \right)^2 / N \; , \tag{2b}$$

where N (*large*) is the total number of measurements made; therefore, we will take $\overline{B''}$ and $\sigma_{B''}^2$ to be *known quantities* in the following analysis.

The threat decision metric of interest to us here is again simply $D'' = S'' - \overline{B''}$, *i.e.* the difference between the total number of counts that our detector records during a time interval Δt and the average background signal that one would expect from a "clean" container during the same interval. The mean and variance of D'' are given by:

$$\overline{D''} = \langle D'' \rangle = \langle S'' - \overline{B''} \rangle = \langle T'' + B'' \rangle - \overline{B''} = \overline{T''} + \overline{B''} - \overline{B''} = \overline{T''} \tag{3a}$$

$$\sigma_{D''}^2 = \sigma_{S''}^2 + \sigma_{\overline{B''}}^2 = \sigma_{T''}^2 + \sigma_{B''}^2 + \sigma_{\overline{B''}}^2 = \overline{T''} + \sigma_{B''}^2 + \sigma_{\overline{B''}}^2 \rightarrow \overline{T''} + \sigma_{B''}^2 \; , \tag{3b}$$

where we have made use of the fact that, for *N* large (*i.e.* after we have inspected many "clean" containers), $\sigma_{\overline{B''}}^2 = \sigma_{B''}^2 / N \rightarrow 0$ (*i.e.* the metric becomes more robust). The process of deciding whether or not a given measured value of D'' implies the presence of a real threat (*i.e.* $T'' > 0$)

involves determining whether or not it can be confidently identified as a "statistical outlier" in relation to the known distribution of data on "clean" containers. In our case, this essentially reduces to determining the "statistical significance" of D'' based on the known (or hypothesized) values of $\langle D'' \rangle$ and $\sigma_{D''}^2$ [5]. Having assumed Gaussian-like distributions, we may therefore define a standard normal test statistic of the form:

$$z = \frac{D'' - \langle D'' \rangle}{\sigma_{D''}} \tag{4}$$

to help us to make quantitative decisions of this sort about the value of D''.

In spite of the *potentially* dire consequences of failing to detect a hidden nuclear threat in a cargo container (*e.g.* admitting entry of clandestine nuclear material or even a nuclear device), the intense nature of operations at a port facility will usually require that containers be given a "speedy trial" (*e.g.* $\approx$ 1-2 min interrogation times) during which they are presumed to be "innocent until proven guilty beyond a reasonable doubt" in order to avoid (or at least minimize) the almost *certain* drawbacks associated with generating an excessive number of false alarms (*e.g.* impeding the flow of commerce, developing a reputation as "the system that cries wolf", *etc.*); therefore, we will adopt $\langle D'' \rangle = \overline{T}'' = 0$ ($\Rightarrow \sigma_{D''}^2 = \sigma_{B''}^2$) as our "null hypothesis" and recast the test statistic of Eqn. (4) as:

$$z_1 = z(\overline{T}'' = 0) = \frac{D''}{\sigma_{B''}} \tag{5}$$

We may then specify a critical value of D'' via:

$$D''(crit) = z_1(crit)\sigma_{B''} \tag{6}$$

which makes P_{fa} (false alarm) $= P(z_1 \geq z_1(crit)) = \alpha$ is as small as needed to meet operational goals (*e.g.* $z_1(crit) = +3.100 \Rightarrow \alpha = 0.001$ (based on a standard normal distribution))[2]. If $D'' \geq D''(crit)$, we may confidently reject our "null hypothesis" and declare the container to be a nuclear threat knowing that the probability of our being *wrong* is less than or equal to α. Sample plots of $D''(crit)$ vs. $\sigma_{B''}$ and $\overline{B}''$ are shown in Figures 1a and 1b, respectively.

Once the critical value of D'' has been specified for a given acceptable probability of false alarm $P_{fa} = \alpha$ and known value of $\sigma_{B''}$, we may assess the "power" $\pi(\overline{T}'' | \alpha)$ of the decision metric (*i.e.* the probability that we can detect an average threat signal $\overline{T}'' > 0$) by recasting the test statistic of Eqn. (4) as:

$$z_2 = z(\overline{T}'' > 0) = \frac{D''(crit) - \overline{T}''}{\sqrt{\overline{T}'' + \sigma_{B''}^2}} = \frac{z_1(crit)\sigma_{B''} - \overline{T}''}{\sqrt{\overline{T}'' + \sigma_{B''}^2}} \tag{7}$$

---

[2] Note that standard normal probabilities such as $P(z \geq z(crit)) = 1 - P(z \leq z(crit))$ may be calculated in spreadsheet programs such as MS Excel® using the *NORMDIST()* function (*e.g.* $P(z \geq z(crit)) = (1-NORMDIST(z(crit)))$.

and then plotting $\pi(\overline{T}''|\alpha) = P(z_2 \geq z_2(z_1(crit), \overline{T}'')) = P_d$ (detection) as a function of $\overline{T}''$ (*cf.* Figure 2 for sample plots of $\pi(\overline{T}''|\alpha)$ vs. $\overline{T}''$ at fixed values of $\sigma_{B''}$). Alternatively, if we have already defined a specific target value of $P_d = P(z_2 \geq z_2(crit)) = \pi(\overline{T}''(crit)|\alpha)$ based on performance requirements (*e.g.* $\pi(\overline{T}''(crit)|\alpha) = 0.950 \Rightarrow z_2(crit) = -1.645$), we may solve Eqn. (7) directly to obtain $\overline{T}''(crit)$ as a function of $\sigma_{B''}$:

$$\overline{T}''(crit) = \frac{1}{2}\left\{ 2z_1(crit)\sigma_{B''} + z_2(crit)^2 - z_2(crit)\sqrt{4\sigma_{B''}^2 + 4z_1(crit)\sigma_{B''} + z_2(crit)^2} \right\} . \quad (8)$$

Sample plots of $\overline{T}''(crit)$ vs. $\sigma_{B''}$ and $\overline{B}''$ are shown in Figures 3a and 3b, respectively.

The signal-to-noise ratio (SNR) associated with a specific value of $\overline{T}''(crit)$ may also be of interest from an experimental point of view. By definition, this is given by:

$$SNR(\overline{T}''(crit)) = \frac{\overline{T}''(crit)}{\sqrt{\sigma_{T''(crit)}^2 + \sigma_{B''}^2}} = \frac{\overline{T}''(crit)}{\sqrt{\overline{T}''(crit) + \sigma_{B''}^2}} , \quad (9)$$

where $\overline{T}''(crit)$ has the functional form shown in Eqn. (8). Sample plots of $SNR(\overline{T}''(crit))$ vs. $\sigma_{B''}$ and $\overline{B}''$ are shown in Figures 4a and 4b, respectively. It is interesting to note that $SNR(\overline{T}''(crit))$ $\rightarrow z_1(crit) - z_2(crit)$ as $\sigma_{B''} \rightarrow \infty$; therefore, for a plausible set of NCW operational parameters such as $P_{fa} = P(z_1 \geq z_1(crit)) = \alpha = 0.001$ and $P_d = P(z_2 \geq z_2(crit)) = \pi(\overline{T}''(crit)|\alpha) = 0.950$, $SNR(\overline{T}''(crit))$ (*i.e.* the threshold SNR for $\overline{T}''$) should not exceed a value of 4.745, regardless of the measured value of $\sigma_{B''}$.

Finally, so-called "Receiver Operating Characteristic" (ROC) curves are generally accepted as the standard way to graphically summarize the performance of a decision metric. In our case, ROC curves for any specific value of $\overline{T}''$ may be obtained by plotting $P_d = P(z \geq z_2(z_1, \overline{T}''))$ vs. $P_{fa} = P(z \geq z_1)$, where $z_2(z_1, \overline{T}'')$ has the basic functional form shown in Eqn. (7). Sample ROC curves at a fixed value of $\sigma_{B''}$ are shown in Figure 5.

3. Extrapolation to particles at detector interface

Unless the detectors associated with our inspection system are 100% efficient (not likely to be the case), the next step in the analysis of our threat decision metric will be to extrapolate the various critical values defined in § 2 above in terms of measured counts (*e.g.* $D''(crit)$, $\overline{T}''(crit)$ and $SNR(\overline{T}''(crit))$) back to their counterparts defined in terms of particles arriving at the detector interface (*e.g.* $D'(crit)$, $\overline{T}'(crit)$ and $SNR(\overline{T}'(crit))$). Although we rarely stop to think about it, detector counting statistics actually reflect the conditional binomial probability $h(y, \xi|x)$ of recording exactly y out of a possible x incident particles given a fixed "probability of success" (*i.e.* detection efficiency) ξ for each particle (*cf.* Appendix A). With this in mind, it can readily be shown that $D'(crit)$ and $\overline{T}'(crit)$ are given by:

$$D'(crit) = D''(crit)/\xi \quad (10)$$

$$\overline{T}'(crit) = \overline{T}''(crit)/\xi , \quad (11)$$

which certainly comes as no surprise; however, the general relationship between $SNR(\overline{T}'(crit))$ and $SNR(\overline{T}''(crit))$ is not quite as straightforward:

$$SNR(\overline{T}'(crit)) = \frac{\overline{T}'(crit)}{\sqrt{\sigma_{T'}^2 + \sigma_{B'}^2}} = \frac{\overline{T}'(crit)}{\sqrt{\overline{T}'(crit) + \sigma_{B'}^2}}$$

$$= \frac{\overline{T}''(crit)\big/\xi}{\sqrt{\overline{T}''(crit)\big/\xi + \left(\sigma_{B''}^2 - (1-\xi)\,\overline{B}''\right)\big/\xi^2}} \quad . \tag{12a}$$

If (again, as is likely the case) the background signal B'' is distributed Poisson (*i.e.* $\sigma_{B''}^2 = \overline{B}''$), then Eqn. (12a) reduces to the familiar result:

$$SNR(\overline{T}'(crit)) = \frac{\overline{T}''(crit)\big/\xi}{\sqrt{\overline{T}''(crit)\big/\xi + \overline{B}''\big/\xi}} = \frac{\overline{T}''(crit)\big/\xi}{\sqrt{\left(\overline{T}''(crit) + \overline{B}''\right)\big/\xi}}$$

$$= \frac{\overline{T}''(crit)\big/\xi}{\sqrt{\left(\overline{T}''(crit) + \sigma_{B''}^2\right)\big/\xi}} = SNR(\overline{T}''(crit))\big/\sqrt{\xi} \quad . \tag{12b}$$

4. Estimating true threat signal strength

Having defined the critical SNM threat signal level for our decision metric in terms of both measured counts ($\overline{T}''(crit)$) and particles arriving at the detector interface ($\overline{T}'(crit)$) for a given set of system operational parameters (*e.g.* P_{fa} and P_d), the operative question becomes, "What does this tell us about the critical value of the *true* threat signal $\overline{T}(crit)$ at its origin and, *in particular*, what does it tell us about the smallest *mass* of SNM that we could reasonably expect to detect with a cargo inspection system such as the NCW?" The answer in both cases is, unfortunately, almost nothing (how anticlimactic!).

The critical value of the true SNM threat signal at its origin will depend on the detailed design of the inspection system *and* the specific threat scenario considered. While the background signal statistics (*e.g.* $\overline{B}''$ and $\sigma_{B''}^2$) can, in principal, be measured with almost arbitrary precision by averaging over many "clean" containers, the exact *values* of $\overline{B}''$ and $\sigma_{B''}^2$ (and hence the values of $\overline{T}''(crit)$ and $\overline{T}'(crit)$ for our decision metric) will depend on the efficiency of the detectors used in the inspection system and their location relative to the container. Even if the background signal statistics are assumed to be well known, the true SNM threat signal will still depend on a *plethora* of other factors including the intensity and energy spectrum of the neutron probe used by the inspection system, the elemental composition and physical arrangement of legitimate cargo in the container, the physical configuration of the SNM (*e.g.* spheres, cylinders, plates, powder, *etc.*) and its proximity to legitimate (or clandestine) materials which might accidentally (or intentionally) tend to conceal its presence in some way and, needless to say, the actual location of the SNM in relation to the neutron source and detector system (note that all factors associated with the shape of the SNM, its shielding (whether accidental *or* intentional) and its final placement in the container are at the complete discretion of the nefarious characters at-

tempting to smuggle the material in). Although some simplifying assumptions can certainly be made (*e.g.* homogeneous cargo distributions, pure exponential attenuation of neutron and γ-ray fluxes, *etc.*), the shear complexity of this eclectic mix of largely indeterminate variables makes the task of extracting a unique value for the true SNM threat signal $\overline{T}(crit)$ at its origin analytically intractable from any practical point of view; however, as intrepid scientists we needn't be discouraged by this minor inconvenience, because this is a perfect example the sort of problem that God created Monte Carlo codes to address!

Modern, general-purpose Monte Carlo radiation transport algorithms such as LLNL's COG code package [6] are now capable of simulating not only the physical layout of proposed cargo inspection systems such as the NCW, but also the detailed characteristics of the system's high-intensity $D(d,n)^3He$ neutron source [7] and the β-delayed γ-ray spectra emitted during the decay of neutron-induced SNM fission products [8]. A series of Monte Carlo simulations of the NCW system could easily be done with different physical configurations of SNM placed in the most challenging locations in a container with "worst-probable-case" cargo configurations to obtain credible values for the limiting number of γ rays that one might expect to see at the detector interface in those scenarios. When coupled with measured detector efficiencies and either measured values or "worst-case" simulations of the background signal, these calculations could then be used to at least place an upper bound on the value of $\overline{T}(crit)$ and thereby derive an estimate for the smallest mass of SNM that we could reliably detect.

5. Summary and comments

The statistical properties of the threat decision metric $D'' = S'' - \overline{B}''$ have been described and procedures have been outlined for using it in conjunction with a cargo inspection system such as the NCW to define critical (*i.e.* minimum detectable) SNM threat signal levels in terms of both measured counts ($\overline{T}''(crit)$) and particles arriving at the detector interface ($\overline{T}'(crit)$) for a given set of system operational parameters (*e.g.* P_{fa} and P_d). It was also noted that that, regardless of how well the background signal statistics (*e.g.* $\overline{B}''$ and $\sigma^2_{B''}$) are eventually known, the complex mix of largely indeterminate variables associated with NCW inspection scenarios will make the task of extracting a unique value for the *true* SNM threat signal $\overline{T}(crit)$ at its origin analytically intractable; however, detailed Monte Carlo simulations of a carefully-chosen set of "worst-case" inspection scenarios should allow us to at least bound the problem and derive meaningful estimates for $\overline{T}(crit)$ and the smallest detectable mass of SNM.

Finally, it should be noted that, while the present report has focused on the potential use of this decision metric with the NCW inspection system, it is perhaps even better suited for use in conjunction with the so-called "long dwell time" problem in which small radiation detectors are placed in individual containers aboard ship and allowed to integrate over the local background and any potential SNM threat signal for the entire time that the ship is at sea in transit between ports. In this case, the passive SNM threat signal (if any) will be due to β-delayed γ rays from spontaneous fission products along with α decay of certain SNM isotopes and the background signal will be due almost entirely to cosmic rays and the decay of naturally-occurring radioisotopes in the surrounding cargo. Since these radiation sources are known to be distributed Poisson, the procedures for using the metric would actually be more straightforward.

Acknowledgements

References

[1] D. Slaughter, M. Accatino, A. Bernstein, J. Candy, A. Dougan, J. Hall, A. Loshak, D. Manatt, A. Meyer, B. Pohl, S. Prussin, R. Walling and D. Weirup, "Detection of special nuclear material in cargo containers using neutron interrogation", LLNL Internal Report UCRL-ID-155315 (2003), available from DOE via: http://www.osti.gov/bridge; see also D. Slaughter, *et al.*, "Preliminary results utilizing high-energy fission product γ rays to detect fissionable material in cargo", Nuc. Inst. & Meth., B241, 777 (2005).

[2] E. Norman, S. Prussin, R.-M. Larimer, H. Shugart, E. Browne, A. Smith, R. McDonald, H. Nitsche, P. Gupta, M. Frank and T. Gosnell, "Signatures of special nuclear material: high-energy γ rays following fission", Nuc. Inst. & Meth., A521, 608 (2004).

[3] J. Pruet, S. Prussin, M.-A. Descalle and J. Hall, "Spectra observed following cargo interrogation", LLNL Internal Report UCRL-PROC-217175 (2005), also published in proceedings of IEEE Nuclear Science Symposium (NSS), San Juan, PR, October 24-27, 2005.

[4] D. Slaughter (Principal Investigator for NCW project), LLNL, private communication.

[5] V. Barnett and T. Lewis, *Outliers in statistical data*, 3rd ed., New York, Wiley, 1994.

[6] R. Buck and E. Lent, "COG: A new, high-resolution code for modeling radiation transport", published in *Energy & Technology Review*, LLNL, June 1993; see also R. Buck, *et al.*, *COG User's Manual*, 5th ed., LLNL, September 2002.

[7] User-defined COG source model for $D(d,n)^3He$ reactions written by J. Hall, LLNL.

[8] J. Hall, J. Pruet, D. Brown, M.-A. Descalle and G. Hedstrom, "Modeling the production of beta-delayed gamma rays for the detection of special nuclear material", LLNL Internal Report UCRL-TR-209738 (2005), available from DOE via: http://www.osti.gov/bridge.

Appendix A (a brief statistics lesson)

Let X be a discrete random variable (RV) with probability distribution function (PDF) $g(x)$ that represents the total number of events associated with some physical process of interest (we will avoid making assumptions about $g(x)$ here - perhaps it is Poisson, perhaps not) and let $Y \leq X$ be a related discrete RV with conditional PDF $h(y, \xi \mid x)$ that represents the subset of these events that arrive at some remote location or are recorded by a physical detector of some sort, in either case with a probability of success ξ ($h(y, \xi \mid x)$ is simply the binomial probability of recording exactly y out of a possible x events given a fixed value of ξ). By definition then, the expected value of any function of Y may be written as:

$$\langle f(Y) \rangle = \sum_{x=0}^{X} \left\{ \sum_{y=0}^{x} f(y) h(y, \xi \mid x) \right\} g(x)$$

$$= \sum_{x=0}^{X} \left\{ \sum_{y=0}^{x} f(y) \left\{ \frac{x!}{y!(x-y)!} \right\} \xi^{y} (1-\xi)^{x-y} \right\} g(x) . \tag{A1}$$

An expression for the expected (mean) value of Y may be derived by setting $f(Y) = Y$ in Eqn. (A1) (note that the first term in the inner sum over y may be dropped in this case since it does not contribute). Thus:

$$\overline{Y} = \langle Y \rangle = \sum_{x=0}^{X} \left\{ \sum_{y=1}^{x} y \left\{ \frac{x!}{y!(x-y)!} \right\} \xi^{y} (1-\xi)^{x-y} \right\} g(x)$$

$$= \sum_{x=0}^{X} \left\{ \sum_{y=1}^{x} \left\{ \frac{x!}{(y-1)!(x-y)!} \right\} \xi^{y} (1-\xi)^{x-y} \right\} g(x)$$

$$= \xi \sum_{x=0}^{X} x \left\{ \sum_{y=1}^{x} \left\{ \frac{(x-1)!}{(y-1)!(x-y)!} \right\} \xi^{y-1} (1-\xi)^{x-y} \right\} g(x)$$

$$= \xi \sum_{x=0}^{X} x \left\{ \sum_{v=0}^{x-1} \left\{ \frac{(x-1)!}{v!((x-1)-v)!} \right\} \xi^{v} (1-\xi)^{(x-1)-v} \right\} g(x) \qquad (\dagger)$$

$$= \xi \sum_{x=0}^{X} x \{1\} g(x) = \xi \langle X \rangle = \xi \overline{X} , \tag{A2}$$

where we replaced $y-1$ with the dummy variable v at ($\dagger$) and recognized that the sum over v then included all possible outcomes of a binomial PDF based on $x-1$ trials and was, therefore, equal to 1. While this was instructive, the fact that $\overline{Y} = \xi \overline{X}$ certainly comes as no surprise. Be patient, though – the relationship between σ_Y^2 and σ_X^2 is more interesting.

The easiest way to derive an expression for σ_Y^2 at this point is to make use of the following identities (true for any random variable V):

$$\sigma_V^2 = \langle V^2 \rangle - \langle V \rangle^2 = \langle V(V-1) \rangle + \langle V \rangle - \langle V \rangle^2 . \tag{A3}$$

Since we already know that $\langle Y \rangle = \xi \langle X \rangle$, the only other quantity we need to find is $\langle Y(Y-1) \rangle$. Following the same basic procedure used in Eqn. (A2) (and noting that this time the first *two* terms may be dropped from the sum over y since neither contributes), it can readily be shown that $\langle Y(Y-1) \rangle = \xi^2 \langle X(X-1) \rangle$. Thus, σ_Y^2 may be written as:

$$\sigma_Y^2 = \langle Y(Y-1) \rangle + \langle Y \rangle - \langle Y \rangle^2$$

$$= \xi^2 \langle X(X-1) \rangle + \xi \langle X \rangle - \left(\xi \langle X \rangle \right)^2$$

$$= \xi^2 \langle X^2 \rangle - \xi^2 \langle X \rangle + \xi \langle X \rangle - \xi^2 \langle X \rangle^2$$

$$= \xi \langle X \rangle - \xi^2 \langle X \rangle^2 + \xi^2 \left(\langle X^2 \rangle - \langle X \rangle^2 \right)$$

$$= \xi(1-\xi)\langle X \rangle + \xi^2 \sigma_X^2 = \xi(1-\xi)\overline{X} + \xi^2 \sigma_X^2 . \tag{A4}$$

The result for σ_Y^2 derived in Eqn. (A4) is intuitively reasonable since we expect the variance in Y to be related to the variance due to attenuation for a given value of X *plus* the variance in X itself. Since a given value of X produces Y events with binomial variance $\xi(1-\xi)$, the total variance due to attenuation is just $\xi(1-\xi)\overline{X}$. Similarly, the variance in X (σ_X^2) is subject to an average gain of ξ, so this component of the variance in Y is weighted by ξ^2. Note that, if X is Poisson (*i.e.* if $\sigma_X^2 = \overline{X}$), then Eqn. (A4) reduces to the familiar result $\sigma_Y^2 = \xi \overline{X} = \overline{Y}$.

Equation (A4) can also be inverted with the help of Eqn. (A2) to derive an expression for σ_X^2 in terms of σ_Y^2 and $\overline{Y}$ if desired:
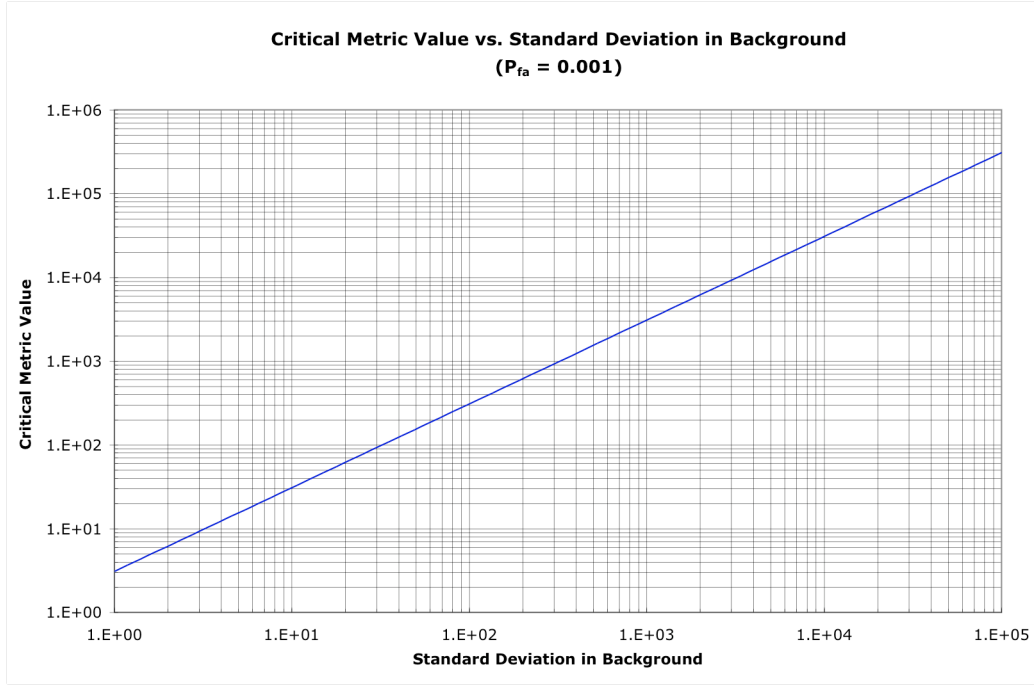
$$\sigma_X^2 = \left(\sigma_Y^2 - (1-\xi)\overline{Y} \right) / \xi^2 \tag{A5}$$

(note that, as with Eqn. (A4), if Y is Poisson (*i.e.* if $\sigma_Y^2 = \overline{Y}$), then Eqn. (A5) reduces to the familiar result $\sigma_X^2 = \overline{Y}/\xi = \overline{X}$). Thus, if σ_X^2 is unknown and if σ_Y^2 can be estimated from experimental data using a standard (unbiased) estimator of the form:

$$\hat{\sigma}_Y^2 = \sum_{i=1}^{N} \left(Y_i - \overline{Y} \right)^2 / (N-1) , \tag{A6}$$

then an unbiased estimate for σ_X^2 may be obtained by simply calculating:

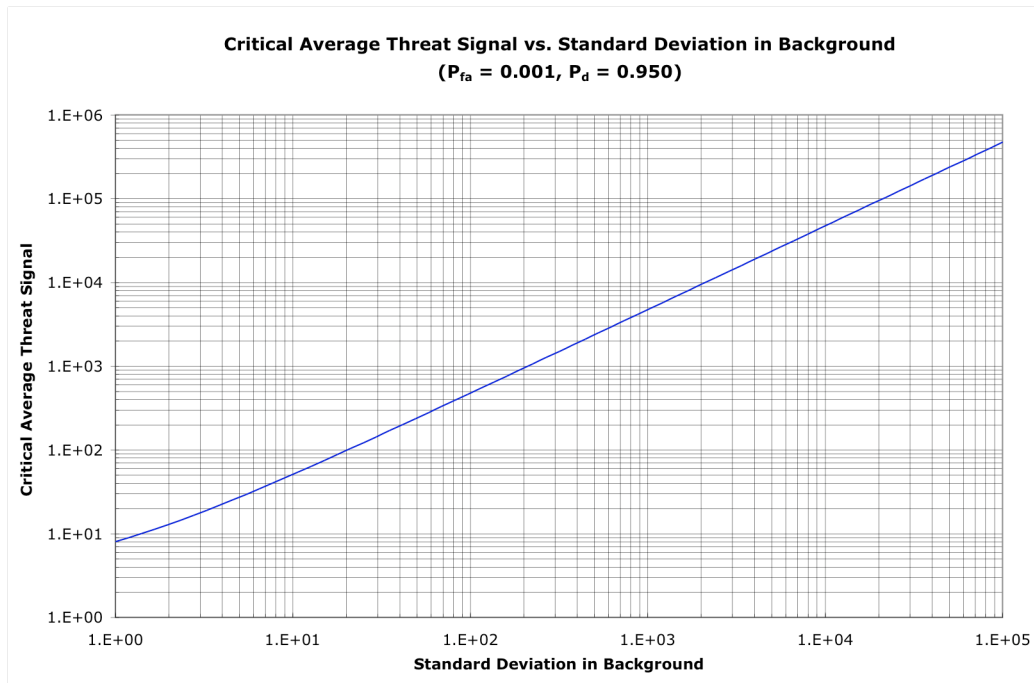$$\hat{\sigma}_X^2 = \left(\hat{\sigma}_Y^2 - (1-\xi)\overline{Y} \right) / \xi^2 . \tag{A7}$$

Figure 1a: Sample plot of critical metric value (measured counts) vs. standard deviation in background signal for P_{fa} = 0.001. If the measured value of D'' exceeds $D''(crit)$, then we may confidently declare the container to be a nuclear threat knowing that the probability of our being *wrong* is less than or equal to P_{fa}.
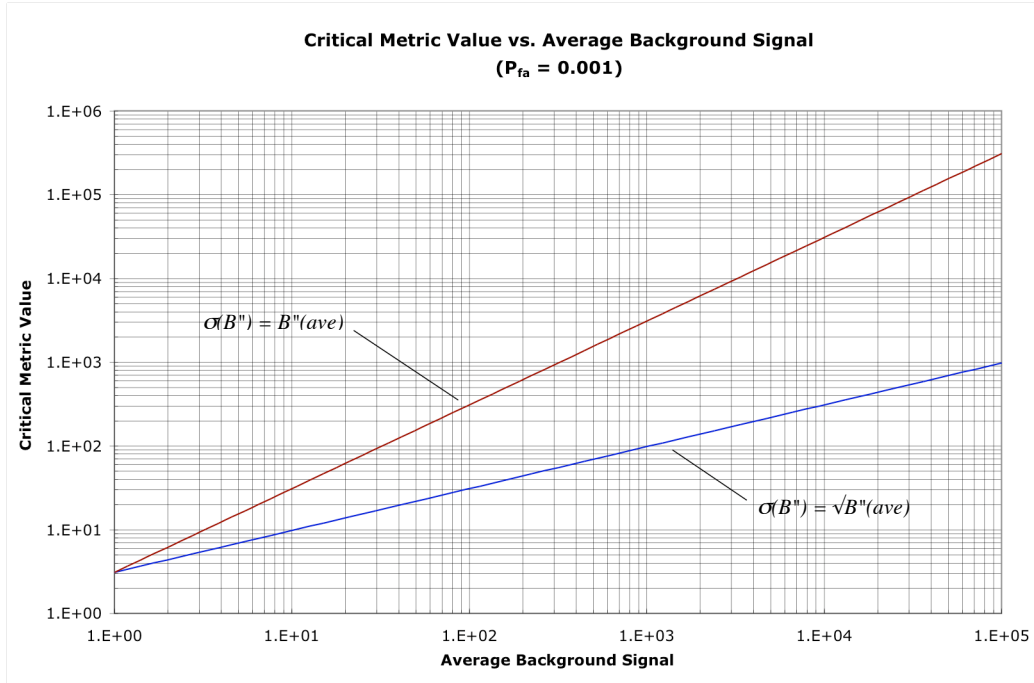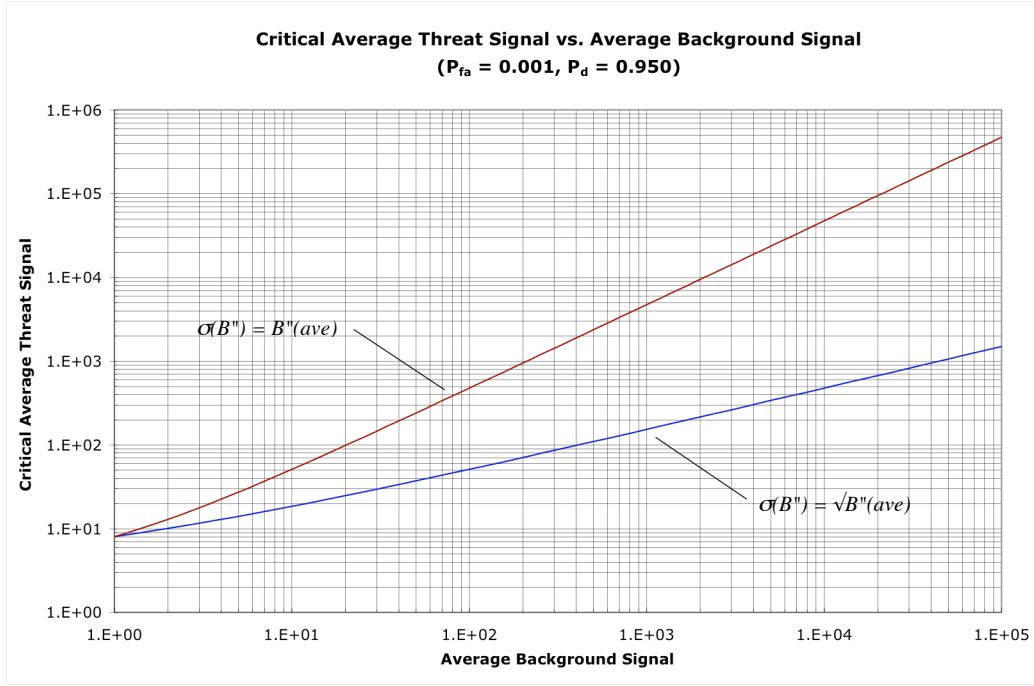

Figure 1b: Sample plots of critical metric value (measured counts) vs. average background signal for P_{fa} = 0.001. The blue curve represents the case in which the background signal is distributed Poisson (*i.e.* $Var(B'') = B''(ave)$) while the red curve represents the pessimistic case in which $Var(B'') = B''(ave)^2$.
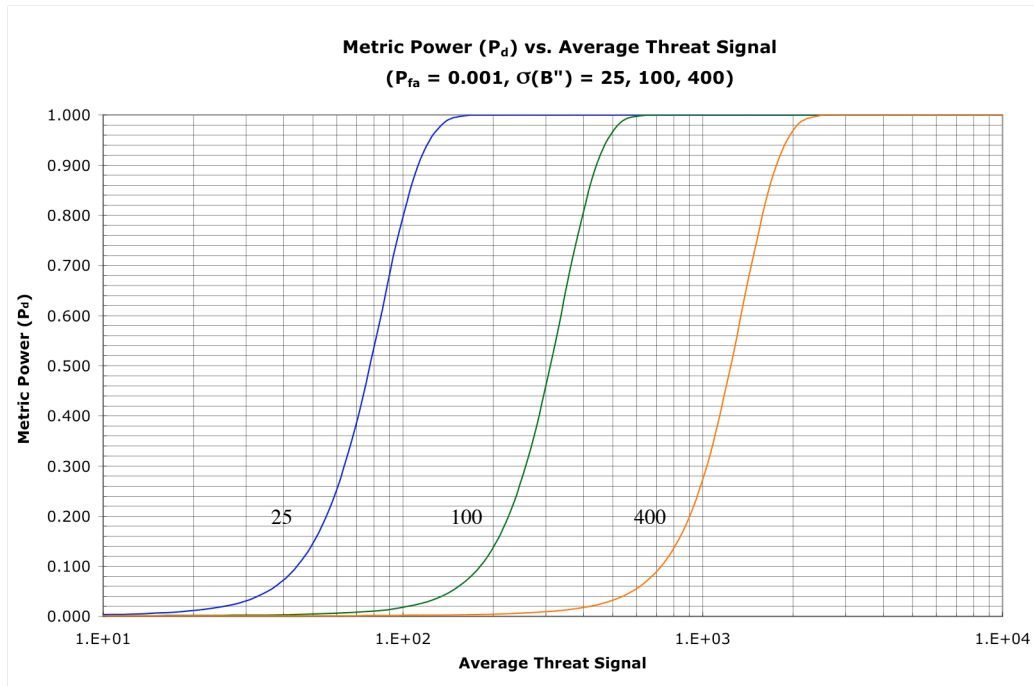
Figure 2: Sample plots of metric power (*i.e.* probability of detection P_d) vs. average threat signal (measured counts) for $P_{fa} = 0.001$ and standard deviations in the background signal of $\sigma(B'') = 25$, 100 and 400. This type of plot is useful in assessing the general relationship between threat signal levels and P_d.
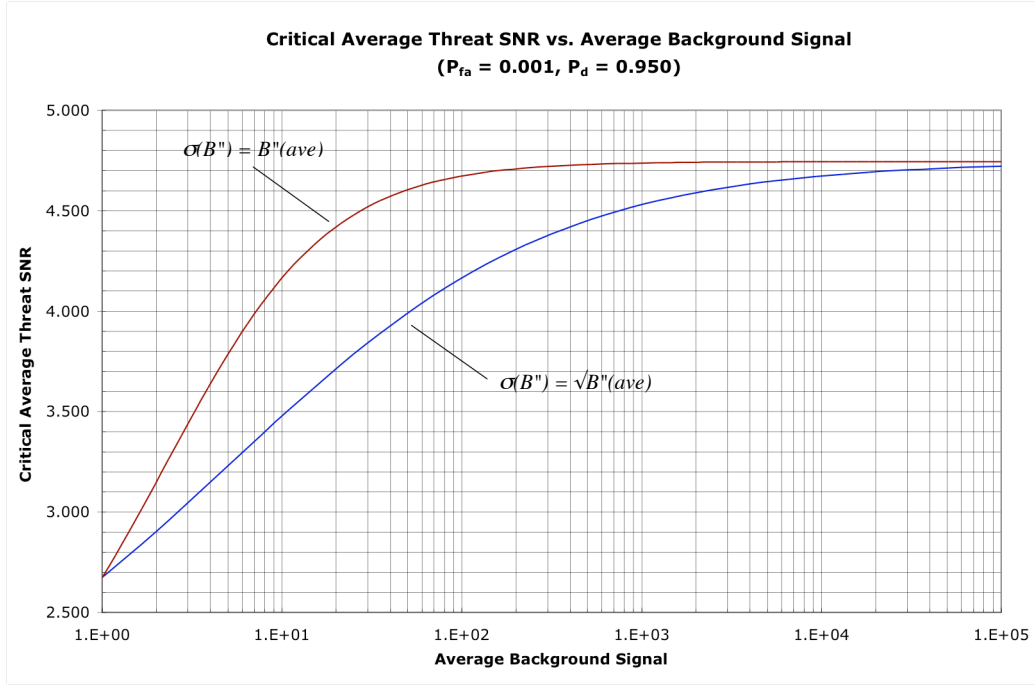

Figure 3a: Sample plot of critical average threat signal (measured counts) vs. standard deviation in background signal for $P_{fa} = 0.001$ and $P_d = 0.950$. This type of plot is useful in determining the critical threat signal if a specific target value of P_d has already been defined based on performance requirements.

Figure 3b: Sample plots of critical average threat signal (measured counts) vs. average background signal for $P_{fa} = 0.001$ and $P_d = 0.950$. The blue curve represents the case in which the background signal is distributed Poisson (*i.e.* $Var(B") = B"(ave)$) while the red curve represents the pessimistic case in which $Var(B") = B"(ave)^2$.
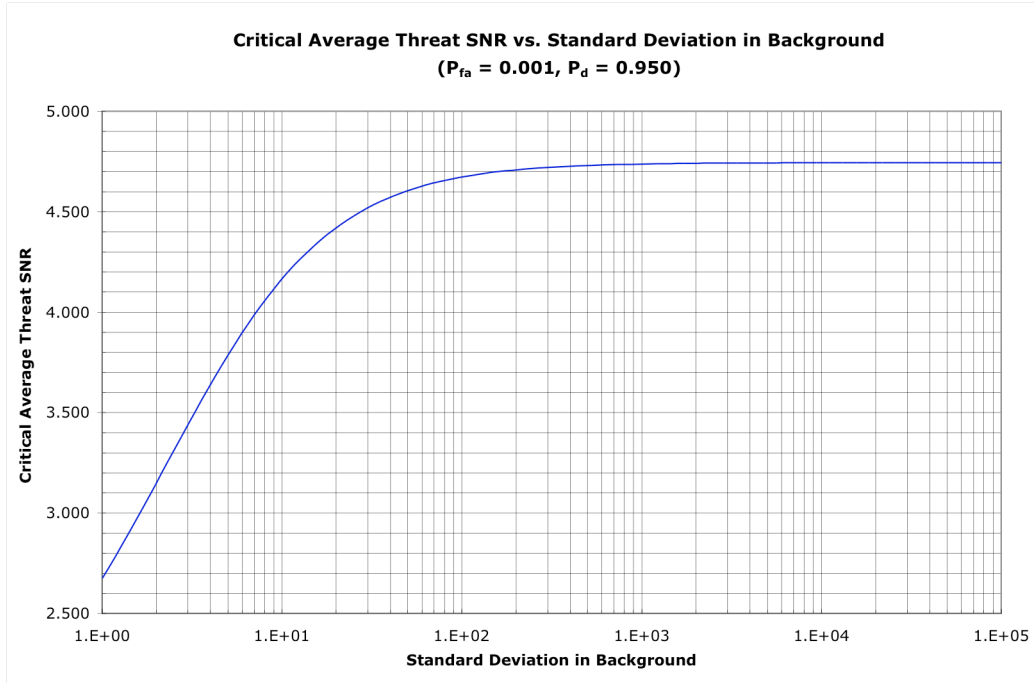
Figure 4a: Sample plot of critical average threat SNR (measured counts) vs. standard deviation in background signal for $P_{fa} = 0.001$ and $P_d = 0.950$. Note that the curve asymptotically approaches $z_1(crit)-z_2(crit) = 4.745$ as $\sigma(B") \rightarrow \infty$; therefore, an experimental SNR ≥ 4.745 should always provide $P_{fa} = 0.001$ and $P_d = 0.950$.

Figure 4b: Sample plots of critical average threat SNR (measured counts) vs. average background signal for $P_{fa} = 0.001$ and $P_d = 0.950$. The blue curve represents the case in which the background signal is distributed Poisson (*i.e.* $Var(B'') = B''(ave)$) while the red curve represents the pessimistic case in which $Var(B'') = B''(ave)^2$.
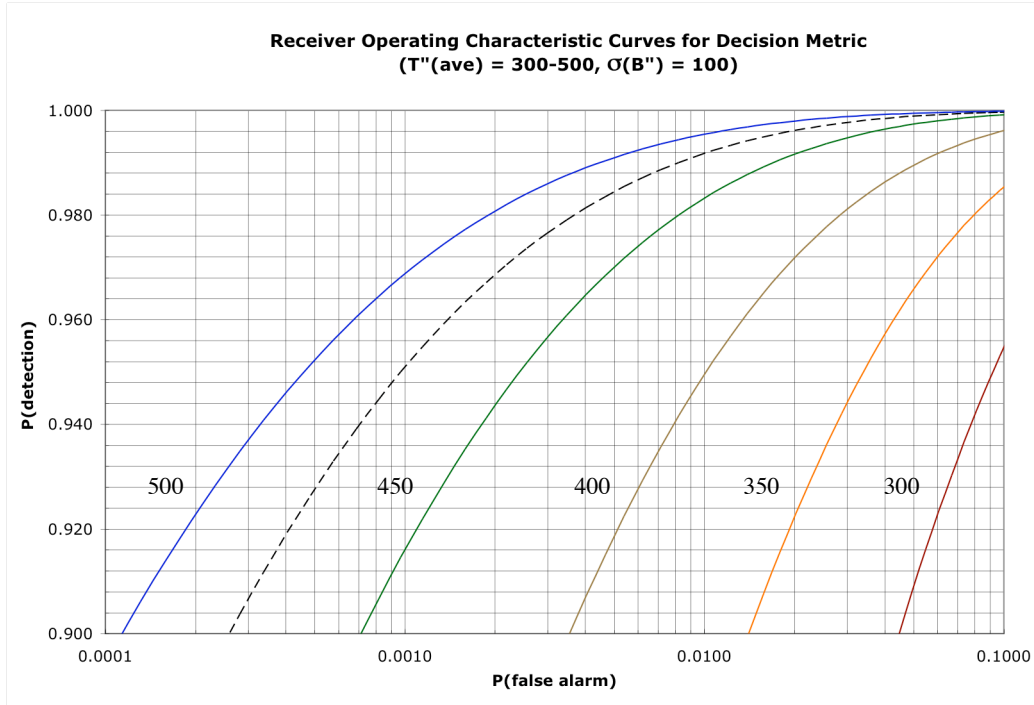
Figure 5: Sample "Receiver Operating Characteristic" (ROC) curves for average threat signal levels $T''(ave) = 300$-500 assuming a fixed standard deviation in the background signal of $\sigma(B'') = 100$. The dashed curve represents the critical value of $T''(ave)$ for which it is possible to simultaneously obtain $P_{fa} = 0.001$ and $P_d = 0.950$.